DOCUMENT RESUME

ED 381 985 EC 303 941

AUTHOR	Elliott, Stephen N.
TITLE	Creating Meaningful Performance Assessments: ERIC Digest E531.
INSTITUTION	Council for Exceptional Children, Reston, Va.; ERIC Clearinghouse on Disabilities and Gifted Education, Reston, VA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-EC-94-2
PUB DATE	Jun 95
CONTRACT	RR93002005
NOTE	3p.; Derived from "Creating Meaningful Performance Assessments: Fundamental Concepts," by Stephen N. Elliott; see ED 375 566.
AVAILABLE FROM	Clearinghouse on Disabilities and Gifted Education, Council for Exceptional Children, 1920 Association Dr., Reston, VA 22091-1589 ($1 each, $5 minimum order prepaid).
PUB TYPE	Information Analyses - ERIC Clearinghouse Products (071)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Definitions; Elementary Secondary Education; *Evaluation Methods; Guidelines; *Performance; *Student Evaluation; Test Reliability; Test Validity
IDENTIFIERS	ERIC Digests; *Performance Based Evaluation

ABSTRACT

This digest offers principles of performance assessment as an alternative to norm-referenced tests. The definition of performance assessment developed by the U.S. Congress's Office of Technology and Assessment is given, common features are listed, and the terms "performance" and "authentic" are defined. Suggested guidelines for addressing validity in performance assessments focus on internal characteristics of the assessment, the relationship of the measure to similar measures or future performance, and the intended effects of using the instrument. In providing evidence for the reliability and validity of performance assessment, evaluators are urged to address: (1) assessment as a curriculum event; (2) task content alignment with curriculum; (3) scoring and subsequent communications with consumers; and (4) linking and comparing results over time. Teachers are urged to use performance assessments in ways which will interact with instruction. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ERIC
Clearinghouse on Disabilities and Gifted Education
THE COUNCIL FOR EXCEPTIONAL CHILDREN
1920 Association Drive • Reston, VA 22091-1589
1-800-328-0272 • 703-264-9474 • FAX 703-264-9494

ERIC Digest

June 1995 EDO-EC-94-2

E531

ED 381 985

CREATING MEANINGFUL PERFORMANCE ASSESSMENTS

Stephen N. Elliott

Performance assessment is a viable alternative to norm-referenced tests. Teachers can use performance assessment to obtain a much richer and more complete picture of what students know and are able to do.

Defining Performance Assessment

Defined by the U.S. Congress, Office of Technology Assessment (OTA) (1992), as *testing methods that require students to create an answer or product that demonstrates their knowledge and skills*, performance assessment can take many forms including

- Conducting experiments.
- Writing extended essays.
- Doing mathematical computations.

Performance assessment is best understood as a continuum of assessment formats ranging from the simplest student-constructed responses to comprehensive demonstrations or collections of work over time. Whatever format, common features of performance assessment involve

1. Students' construction rather than selection of a response.
2. Direct observation of student behavior on tasks resembling those commonly required for functioning in the world outside school.
3. Illumination of students' learning and thinking processes along with their answers (OTA, 1992).

Performance assessments measure what is taught in the curriculum. There are two terms that are core to depicting performance assessment:

1. Performance: A student's active generation of a response that is observable either directly or indirectly via a permanent product.
2. Authentic: The nature of the task and context in which the assessment occurs is relevant and represents "real world" problems or issues.

How Do You Address Validity in Performance Assessments?

The validity of an assessment depends on the degree to which the interpretations and uses of assessment results are supported by empirical evidence and logical analysis. According to Baker and her associates (1993), there are five internal characteristics that valid performance assessments should exhibit:

1. Have meaning for students and teachers and motivate high performance.
2. Require the demonstration of complex cognition, applicable to important problem areas.
3. Exemplify current standards of content or subject matter quality.
4. Minimize the effects of ancillary skills that are irrelevant to the focus of assessment.
5. Possess explicit standards for rating or judgment.

When considering the validity of a performance test, it is important to first consider how the test or instrument "behaves" given the content covered. Questions should be asked such as:

- How does this test relate to other measures of a similar construct?
- Can the measure predict future performances?
- Does the assessment adequately cover the content domain?

It is also important to review the intended effects of using the assessment instrument. Questions about the use of a test typically focus on the test's ability to reliably differentiate individuals into groups and guide the methods teachers use to teach the subject matter covered by the test.

A word of caution: Unintended uses of assessments can have precarious effects. To prevent the misuse of assessments, the following questions should be considered:

- Does use of the instrument result in discriminatory practices against various groups of individuals?
- Is it used to evaluate others (e.g., parents or teachers) who are not directly assessed by the test?

EC 303941

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Providing Evidence for the Reliability and Validity of Performance Assessment

The technical qualities and scoring procedures of performance assessments must meet high standards for reliability and validity. To ensure that sufficient evidence exists for a measure, the following four issues should be addressed:

1. *Assessment as a Curriculum Event.* Externally mandated assessments that bear little, if any, resemblance to subject area domain and pedagogy cannot provide a valid or reliable indication of what a student knows and is able to do. The assessment should reflect what is taught and how it is taught.

   Making an assessment a curriculum event means reconceptualizing it as a series of theoretically and practically coherent learning activities that are structured in such a way that they lead to a single predetermined end. When planning for assessment as a curriculum event, the following factors should be considered:

   - The content of the instrument.
   - The length of activities required to complete the assessment.
   - The type of activities required to complete the assessment.
   - The number of items in the assessment instrument.
   - The scoring rubric.

2. *Task Content Alignment with Curriculum.* Content alignment between what is tested and what is taught is essential. What is taught should be linked to valued outcomes for students in the district.

3. *Scoring and Subsequent Communications with Consumers.* In large scale assessment systems, the scoring and interpretation of performance assessment instruments is akin to a criterion-referenced approach to testing. A student's performance is evaluated by a trained rater who compares the student's responses to multitrait descriptions of performances and then gives the student a single number corresponding to the description that best characterizes the performance. Students are compared directly to scoring criteria and only indirectly to each other.

   In the classroom, every student needs feedback when the purpose of performance assessment is diagnosis and monitoring of student progress. Students can be shown how to assess their own performances when:

   - The scoring criteria are well articulated.
   - Teachers are comfortable with having students share in their own evaluation process.

4. *Linking and Comparing Results Over Time.* Linking is a generic term that includes a variety of approaches to making results of one assessment comparable to those of another. Two appropriate and manageable approaches to linking in performance assessment include:
   - Statistical Moderation. This approach is used to compare performances across content areas for groups of students who have taken a test at the same point in time.
   - Social Moderation. This is a judgmental approach that is built on consensus of raters. The comparability of scores assigned depends substantially on the development of consensus among professionals.

How Can Teachers Influence Students' Performances?

Performance assessment is a promising method that is achievable in the classroom. In classrooms, teachers can use data gathered from performance assessment to guide instruction. Performance assessment should interact with instruction that precedes and follows an assessment task.

When using performance assessments, students' performances can be positively influenced by:

1. Selecting assessment tasks that are clearly aligned or connected to what has been taught.
2. Sharing the scoring criteria for the assessment task with students prior to working on the task.
3. Providing students with clear statements of standards and/or several models of acceptable performances before they attempt a task.
4. Encouraging students to complete self-assessments of their performances.
5. Interpreting students' performances by comparing them to standards that are developmentally appropriate, as well as to other students' performances.

References

Baker, E. L., O'Neill, H. F., Jr., & Linn, R. L. (1993). Policy and validity prospects for performance-based assessments. *American Psychologist, 48*, 1210-1218.

U.S. Congress, Office of Technology Assessment. (1992, February). *Testing in American schools: Asking the right questions.* (OTA-SET-519). Washington, DC: U.S. Government Printing Office.

Derived from Elliot, S. N. (1994). *Creating Meaningful Performance Assessments: Fundamental Concepts.* Reston, VA: The Council for Exceptional Children. Product # P5059.

ERIC Digests are in the public domain and may be freely reproduced and disseminated.

This publication was prepared with funding from the National Library of Education (NLE), Office of Educational Research and Improvement, U.S. Department of Education, under contract no. RR93002005. The opinions expressed in this report do not necessarily reflect the positions or policies of NLE, OERI, or the Department of Education.